International Journal of Mathematics and Computer Applications Research (IJMCAR)
ISSN(P): 2249-6955; ISSN(E): 2249-8060
Vol. 3, Issue 5, Dec 2013, 29-36

COMBINED APPROACH TO PROPAGATE ALEATORY AND EPISTEMIC UNCERTAINTY IN RISK ASSESSMENT

PALASH DUTTA

Department of Mathematics, Dibrugarh University, Dibrugarh, Assam, India

ABSTRACT

In risk assessment, it is most important to know the nature of all available information, data or model parameters. More often, it is seen that available information is interpreted in probabilistic sense because probability theory is a very strong and well established mathematical tool to deal with aleatory uncertainty. However, not all available information, data or model parameters are affected by aleatory uncertainty and can be handled by traditional probability theory. Imprecision may occur due to scarce or incomplete information or data, measurement error or data obtain from expert judgment or subjective interpretation of available data or information. Thus, model parameters, data may be affected by epistemic uncertainty. Fuzzy set theory or possibility theory can be explored to handle this type of uncertainty (Dutta et al. 2012). Sometimes, it is also seen that some model parameters are affected by aleatory uncertainty and some other model parameters are affected by epistemic uncertainty then there is the need for joint propagation of uncertainties. In this paper, an attempt has been made to combine probability distributions, generalized fuzzy numbers and normal fuzzy numbers within the same framework.

KEYWORDS: Aleatory & Epistemic Uncertainty, Fuzzy Set, Generalized Fuzzy Number, Hybrid Method, Risk Assessment

INTRODUCTION

The aspect of uncertainty is an important and integral to any risk assessment process. For any decision making process involving risk the modelling and quantification of the uncertainties is required. The uncertainties are basically two types viz., aleatory and epistemic.Aleatory uncertainty arises due to inherent variability, natural stochasticity, environmental or structural variation across space or through time, manufacturing or genetic heterogeneity among components or individuals, and Varity of others sources of randomness. On the other hand epistemic uncertainty arises due to the insufficient knowledge about the world, which includes small sample sizes, detection limits, imperfections in scientific understanding etc. When some parameters are affected by aleatory uncertainty and other by epistemic uncertainty, how far computation of the risk, then one can either transform all the uncertainties to one type or use some methods for propagation of the uncertainties. Many researchers have studied the issue of combining probabilistic and possibilistic representation of aleatory and epistemic uncertainty respectively within the same computation of risk. For example, Guyonnet et al, (2003), kentel et al (2004) has proposed hybrid method for join handling of probability and possibility distributions. The hybrid method proposed in Guyonnet et al, (2003) combines the random sampling of probability distribution functions (PDFs) with fuzzy interval analysis on the α-cuts. In order to compare random fuzzy set to a tolerance threshold Guyonnet, et al (2003) performed a post-processing of this result. Baudrit et al (2004) laid bare a shortcoming of this post-processing method. Baudrit et al (2004) showed how the theory of evidence, also called theory of Dempster-Shafer (or theory of belief functions; Shafer, 1976) could provide a simple and rigorous answer to the problem of summarizing the results of the hybrid computation for comparison with a tolerance threshold. In the hybrid approach

proposed by Kentel et al (2004) combined utilization of fuzzy and random variables produces membership functions of risk to individuals at different fractiles of risk as well as probability distributions of risk for various alpha-cut levels of the membership function. Li et al. (2007) has proposed different hybrid method for joint handling of probability and possibility distributions. Dutta et al (2012) proposed a hybrid method to deal with both variability and uncertainty within the same framework of computation of risk.

In this paper, an attempt has been made to combine probability distributions, generalized fuzzy numbers and normal fuzzy numbers within the same framework.

BASIC CONCEPT OF FUZZY SETS

Environmental/human health risk assessment is an important aid in any decision-making process in order to minimize the effects of human activities on the environment. Unfortunately, usually environmental data tends to be vague and imprecise, so uncertainty is associated with any study related with these kinds of data. Fuzzy set theory provides a way to characterize the imprecisely defined variables, define relationships between variables based on expert human knowledge and use them to compute results. In this section, some necessary backgrounds and notions (Dutta et al., 2011a) of fuzzy set theory that will be required in the sequel are reviewed.

Definition

Let X be a universal set. Then the fuzzy subset A of X is defined by its membership function $\mu_A : X \to [0,1]$

Which assign a real number $\mu_A(x)$ in the interval [0, 1], to each element $x \in A$, where the value of $\mu_A(x)$ at x shows the grade of membership of x in A.

Definition

Given a fuzzy set A in X and any real number $\alpha \in [0, 1]$. Then the α -cut or α -level or cut worthy set of A, denoted by $^{\alpha}A$ is the crisp set

$$^{\alpha}A = \left\{ x \in X : \mu_A(x) \ge \alpha \right\} \tag{1}$$

The strong α -cut, denoted by $^{\alpha+}A$ is the crisp set

$$^{\alpha}A = \left\{ x \in X : \mu_A(x) > \alpha \right\} \tag{2}$$

Definition

The support of a fuzzy set A defined on X is a crisp set defined as

$$Supp(A) = \left\{ x \in X : \mu_A(x) > 0 \right\} \tag{3}$$

Definition

The height of a fuzzy set A, denoted by h(A) is the largest membership grade obtain by any element in the set and it is denoted as

$$h(A) = \sup_{x \in X} \mu_A(x) \tag{4}$$

Definition

A fuzzy number is a convex normalized fuzzy set of the real line R whose membership function is piecewise continuous.

Definition

A triangular fuzzy number A can be defined as a triplet [a, b, c]. Its membership function is defined as:

$$\mu_A(x) = \begin{cases} \frac{x-a}{b-a}, a \le x \le b \\ \frac{c-x}{c-b}, b \le x \le c \end{cases} \tag{5}$$

Definition

A trapezoidal fuzzy number A can be expressed as [a, b, c, d] and its membership fuzzy number is defined as:

$$\mu_A(x) = \begin{cases} \frac{x-a}{b-a}, a \le x \le b \\ 1, \quad b \le x \le c \\ \frac{d-x}{d-c}, c \le x \le d \end{cases} \tag{6}$$

GENERALIZED FUZZY NUMBERS (GFN)

The membership function of GFN A= [a,b,c,d;w] where $a \le b \le c \le d$, $0 < w \le 1$ is defined as (Chen, 11985, 1999)

$$\mu_A(x) = \begin{cases} 0, x < a \\ w\frac{x-a}{b-a}, a \le x \le b \\ w, \quad b \le x \le c \\ w\frac{x-c}{d-c}, c \le x \le d \\ 0, x > d \end{cases} \tag{7}$$

If w = 1, then GFN A is a normal trapezoidal fuzzy number A = [a, b, c, d]. If a = b and c = d, then A is a crisp interval .If b = c then A is a generalized triangular fuzzy number. If a = b = c = d and w = 1 then A is a real number. Compared to normal fuzzy number the GFN can deal with uncertain information in a more flexible manner because of the parameter w that represent the degree of confidence of opinions of decision maker’s.

PROPOSED HYBRID APPROACH

Consider the model

$$M = g\ (P_1, P_2,..,P_m, Q_1, Q_2,...,Qr, F_1, F_2 ,...,F_n) \tag{8}$$

which is a function of parameters. Suppose $P_1, P_2 \ldots P_m$ are m parameters presented by probabilistic distributions while $Q_1, Q_2,...,Q_r$ are r parameters presented by generalized fuzzy numbers with heights $w_1, w_2,...,w_r$ respectively and $F_1, F_2 \ldots, F_n$ are n parameters presented by normal fuzzy numbers.

The approach is explained below:

- Consider all r generalized fuzzy numbers $Q_1, Q_2, ..., Q_r$ & n normal fuzzy numbers $F_1, F_2, .., F_n$. Then calculate α-cut for each fuzzy number. As height of each generalized fuzzy numbers are different so to perform computation we consider α values lie within the interval $[0, w]$, where $w = \min(w_1, w_2, ..., w_r)$.

First calculate $\alpha(=0)$-cut for each fuzzy number. Then r+n numbers of closed intervals (as α-cut gives closed intervals) i.e., $2(r+n)$ numbers of values will be obtained.

- Generate m number of uniformly distributed random numbers from [0, 1] and perform Monte Carlo simulation to obtain m number of random numbers by sampling probability distribution.
- Assign all m random numbers and $2(r+n)$ values in the model M and calculate $M_1^{inf} = Inf(M)$ and $M_1^{sup} = Sup(M)$.
- Repeat step1 to step3 for 5000 times
- Plot cumulative distribution function of $(M_1^{inf}, M_2^{inf}, \ldots, M_{5000}^{inf})$ and $(M_1^{sup}, M_2^{sup}, \ldots, M_{5000}^{sup})$.
- Calculate α-cuts for other α values (note that greatest α value is w. i.e., $\alpha \in [0, w]$) of each fuzzy number and repeat step 2 to step 5. If proceed in this way a family of cdfs will be obtained

From theses family of cdfs membership functions (generalized fuzzy numbers) at different fractiles can be generated with each of height w.

A CASE STUDY

To demonstrate and make use of the proposed hybrid method a hypothetical case study for cancer risk assessment is presented here. Suppose water became contaminated due to the release of radionuclide to the water. Need to calculate cancer risk for the ingestion pathway.

The risk assessment model due to the ingestion of radionuclides in water as provided by EPA, 2001 is follows

$$\mathbf{Risk} = \frac{\mathbf{C \times IR \times EF \times ED}}{\mathbf{BW \times AT}} \mathbf{\times CSF} \tag{9}$$

Where C is concentration (mg/L), IR is the ingestion rate (L/day), EF is the exposure frequency (days/year), ED is exposure duration (years), BW is the body weight (kg), AT is averaging time (equal to 70 years x 365 days/year), and CSF is the cancer slope or potency factor associated with ingestion $(\text{mg/kg-day})^{-1}$.

Values of the parameters for the calculation of cancer risk are given in the table 1.

Table 1: Parameter Values Used in the Risk Assessment

Parameter	Units	Type of Variable	Value/Distribution
Concentration (C)	mg/L	Probabilistic	Normal(0.15, 0.0005)
Intake rate(IR)	L/day	GFN	[4.75, 5, 5.25; 0.8]
Exposure frequency (EF)	Days/year	Constant	350
Exposure Duration (ED)	Years	Constant	30
Average Time (AT)	Days	Constant	25550
Body Weight (BW)	Kg	Fuzzy	70
Cancer slope factor (CSF)	$(\text{mg/kg-day})^{-1}$	Constant	[0.14, 0.15, 0.16]

The result of the cancer risk assessment due to the ingestion of radionuclides in water using equation (1) is depicted in figure (1). For simple and clear representation here we consider $\alpha = 0.0,\ 0.4,\ 0.8.$

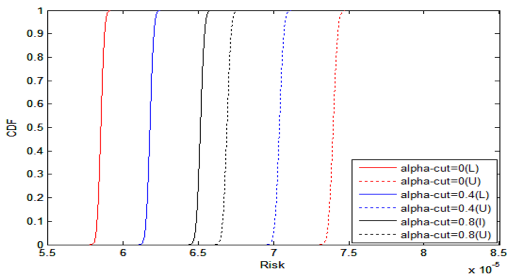

Figure 1: Cumulative Distribution Functions of Cancer Risk When α = 0.0, 0.4, 0.8

In this study, representation of the parameter concentration (*C*) is considered as probabilistic distribution having with mean 0.15 and standard deviation 0.0005, intake rate (*IR*) is taken as generalized fuzzy number while representation of the parameter cancer slope factor is considered as normal triangular fuzzy number and other parameters are considered as constant. Using our proposed method to deal with both type uncertainty natures in the risk assessment we have the result in the form of Cdfs. From these cdfs, risk at different fractiles (Maxwell et al, 1998, Kentel et al, 2004 & Dutta et al, 2011b) can be calculated and which are obtained in the form of generalized fuzzy number. For instance, at 90th fractile, cancer risk value lies in the generalized fuzzy number [5.881e-05, 6.546e-05, 6.721e-05, 7.428e-05; 0.8]. Similarly, at 80th fractile, risk values lie in the generalized fuzzy number [5.872e-05, 6.534e-05, 6.711e-05, 7.417e-05; 0.8]. Their graphical representations are given below.

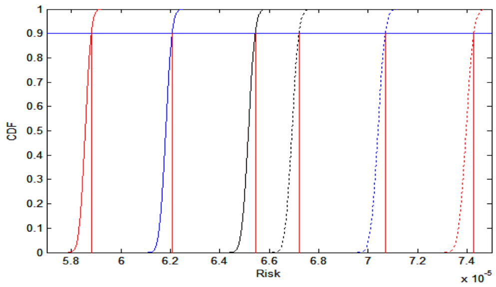

Figure 2: Cumulative Distribution Functions of Cancer Risk When α = 0.0, 0.4, 0.8

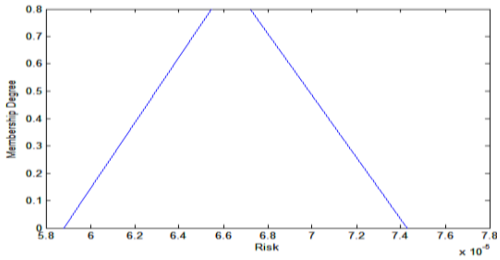

Figure 3: Membership Function of Risk at 90th Fractile

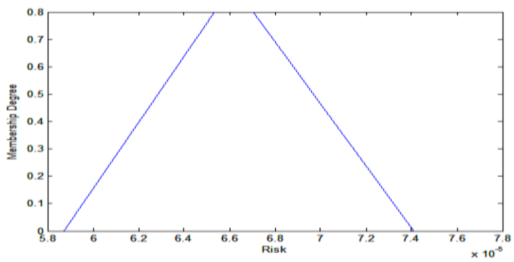

Figure 4: Membership Function of Risk at 80th Fractile

CONCLUSIONS

Risk assessment is an important and popular aid in the decision making process. The aim of risk assessment is to estimate the severity and likelihood of harm to human health from exposure to a substance or activity that under plausible circumstances can cause to human health (Kentel et al. 2004). In risk assessment, it is most important to know the nature of all available information, data or model parameters. More often, it is seen that available information model parameters, data are tainted with aleatory and epistemic uncertainty or both type of uncertainty. When some model parameters are affected by aleatory uncertainty and other some parameters are affected by epistemic uncertainty, how far computation of the risk is concern, one can either transform all the uncertainties to one type of format or need for joint propagation of uncertainties.In this paper, an effort has been made to combine probability distributions, normal fuzzy numbers and generalized fuzzy numbers within the same framework. To demonstrate and make use of the proposed hybrid method a hypothetical case study for cancer risk assessment is carried out here. After performing risk assessment using our approach risk is obtained in the form of Cdfs and from which, membership functions of the risk are generated at different fractiles. The membership functions of risk at different fractiles are generalized fuzzy number since representation of at least one parameter is taken as generalized fuzzy number. Also we have observed that shape of the generalized fuzzy number is trapezoidal type it is because any arithmetic operation of generalized fuzzy numbers (also generalized fuzzy number and normal fuzzy number) produces trapezoidal type generalized fuzzy number.

REFERENCES

1. C. Baudrit, D. Dubois, D. Guyonnet, H. Fargier,Joint Treatment of imprecision and Randomness in Uncertainty Propagation, Proc. Conf. on Information Processing and Management of Uncertainty in Knowledge-Based Systems, Perugia, (2004) 873-880.

2. C. Baudrit, D. Dubois, D. Guyonnet, Joint Propagation and Exploitation of Probabilistic and Possibilistic Information in Risk Assessment, IEEE Transaction on Fuzzy Systems, 14 (2006) 593-608.

3. D. Guyonnet, B. Bourgine, D. Dubois, H. Fargier, B. Côme, J. P. Chilès, Hybrid approach for addressing uncertainty in risk assessments, Journal of Environmental Engineering, 126 (2003) 68-78.

4. E.Kentel, M. M.Aral, Probalistic-fuzzy health risk modeling, Stoch Envir Res and Risk Ass 18(2004) 324-338.

5. J. Li, G. H. Huang, G. M. Zeng, I. Maqsood, Y. F. Huang, An integrated fuzzy-stochastic modeling approach for risk assessment of groundwater contamination, Journal of Environmental Management, 82(2007) 173-188.

6. P. Dutta, H. Boruah T. Ali, Fuzzy Arithmetic with and without using α-cut method: A Comparative Study, International Journal of Latest Trends in Computing.Volume 2, Issue 1 (2011a) pp.: 99-107.

7. P. Dutta, T. Ali, Fuzzy Focal Elements in Dempster-Shafer Theory of Evidence: Case study in Risk Analysis, International Journal of Computer application, Volume 34, No.-1 (2011b), pp.: 46-53.

8. P. Dutta, T. Ali, A Hybrid Method to Deal with Aleatory and Epistemic Uncertainty in Risk Assessment, International Journal of Computer Applications, 42 (2012) 37-44.

9. R. M. Maxwell, S.D. Pelmulder, A. F. B. Tompson, W. E. Kastenberg, On the development of a new methodology for groundwater-driven health risk assessment, Water Resources Res, 34(1998) 833–847.

10. S.H.Chen, (1985) Operations on fuzzy numbers with function principal, *Tamkang J.Manag.Sci.*,vol.6, no.1, pp.13 -25.

11. S.H.Chen, (1999) Ranking generalized fuzzy number with graded mean integration, (in proc.8th Int.Fuzzy Syst.Assoc.World Congr.,vol.2,Taipei,Taiwan,Republic of China, pp.899-902.)

12. U.S. EPA. 2001. Risk Assessment Guidance for Superfund, Volume I: Human Health Evaluation Manual (Part E, Supplemental Guidance for Dermal Risk Assessment). Office of Emergency and Remedial Response, EPA/540/R/99/005, Interim, Review Draft. United States Environmental Protection Agency. September 2001.